

Die Große kommentierte Frankfurter Ausgabe der Werke Thomas Manns

Ein innovatives verlegerisches Konzept

VON INGRID SCHMIDT UND CAROLIN MÜLLER-SPITZER

Alles Gescheite ist schon gedacht worden,
man muss nur versuchen, es noch einmal zu denken.

J. W. von Goethe

1. Einleitung

Wir leben in einer Informationsgesellschaft. Das haben die Menschen zu allen Zeiten getan. Information und die Kommunikation von Wissen waren in jeder Gesellschaftsform wichtig. Wissen ist heute ein Machtfaktor und ist es immer gewesen. Francis Bacon wusste das treffend mit "Knowledge itself is power" zu sagen. Zusammen mit Termini wie Informations- und Wissensmanagement verweisen Informations- oder Wissensgesellschaft heute allerdings in eine Welt, in der diese Begriffe einen zentralen Stellenwert einnehmen; sie betonen ausdrücklich deren Rolle in unserer Gesellschaft.

Wir leben im Computerzeitalter. Das tun die Menschen erst seit ein paar Jahrzehnten und auch nicht überall auf der Welt. Schon vor mehr als 300 Jahren spekulierte der deutsche Philosoph Leibniz über "Rechenapparate zum Finden von Gedanken", die seiner Meinung nach schließlich dazu führen würden, das diskursive Streiten durch ein "Rechnen wir!" zu ersetzen. Damit hatte er unsere heutige Kulturtechnik des 'Computerns' durchaus richtig erahnt.¹ Die Einschätzung des Vorstandsvorsitzenden von IBM 1943 "I think there is a world market for about five computers" und die des Präsidenten von Digital Equipment 1977 "There is no reason for any individual to have a computer in their home" sind da – nach heutiger Sicht – weit weniger zutreffend.

Leibniz sei hier nicht nur wegen seiner Voraussicht unseres Computerzeitalters angeführt, sondern auch im Hinblick darauf, wie man heute innovativ mit verlegerischen Inhalten umgehen kann. Als Bibliothekar ist Leibniz mit Ordnungssystemen von Büchern befasst, muss sie den verschiedenen Wissensgebieten zuordnen und muss sich am Ende immer für einen einzigen

¹ FRANK HARTMANN: Medienphilosophie. Wien 2000, S. 21.

Standort des Buches entscheiden. Im 17. Jahrhundert konkretisieren sich die Pläne für die großen Enzyklopädien, die dann im 18. Jahrhundert erscheinen. Als Gelehrter und Bibliothekar nimmt Leibniz diese Gedanken auf und ist bis zum Ende seines Lebens mit der Idee einer Universalenzyklopädie beschäftigt. Dabei geht es ihm weniger um ein System elementarer Begriffe für eine solche Enzyklopädie, als vielmehr um "ein praktisches und flexibles Instrument, das allen ermöglichen soll, das immense Gebäude des Wissens zu kontrollieren".² Die Organisationsprobleme, die er als Bibliothekar hatte, indem ein Buch nicht eindeutig einem Wissensgebiet zuzuordnen war, gepaart mit seinen Interessen an der 'Wissensorganisation', ließen ihn schließlich folgern, dass

"nichts anderes übrig [bliebe], als eine Enzyklopädie zu organisieren, die wir polydimensional und gemischt nennen würden, eine Enzyklopädie, die [...] eher nach 'Routen' als nach Disziplinen aufgebaut ist, ein Modell des theoretisch-praktischen Wissens, das zu 'transversalen' Benutzungen ermuntert: einmal im theoretischen Sinn nach der Ordnung der Beweise, wie es die Mathematiker tun, das andere Mal analytisch und praktisch, mit Rücksicht auf die menschlichen Ziele; und dann müsste ein Register angeführt werden, das die verschiedenen Themen ebenso aufzufinden erlaubt wie ein gleichbleibendes Thema, das unter verschiedenen Aspekten an verschiedenen Stellen behandelt worden ist".³

Mit diesen Überlegungen hat Leibniz die Gedanken einer dynamischen Organisation von 'Wissen' in semantischen Netzen vorweggenommen, die abhängig von der Fragestellung des Benutzers traversiert werden können. Dabei hebt das vom Computer ermöglichte polydimensionale Ordnungsprinzip die Linearität des Papiers auf.

Heute verhält es sich durchaus so, wie Thomas Mann es sieht, wenn er sagt: "Wissen ist so ziemlich das einzige auf der Welt, was auf der Straße liegt und wovon sich jeder nehmen kann, soviel er Lust hat." Jedoch stößt man gerade bei der freien Verfügbarkeit der heutigen Informationsmengen auf die Notwendigkeit, diese gezielt und ergonomisch filtern zu können. Wie man im Kontext einer literarischen Edition durch den Aufbau eines semantischen Netzes Inhalte zugreifbar machen kann, ist Thema dieses Aufsatzes. Die grundsätzliche Konzeption des Thomas-Mann-Projekts und die Vorüberlegungen, die zu dieser Konzeption führten, haben wir bereits in anderen Publikationen ausführlich beschrieben.⁴ Hier legen wir den Schwerpunkt auf

² UMBERTO ECO: Die Suche nach der vollkommenen Sprache. München 1997, S. 284.

³ Zitiert nach ECO (wie Anm. 2), S. 285.

⁴ INGRID SCHMIDT/CAROLIN MÜLLER: Planning a new type of literary edition: The Thomas Mann Project. In: XML Europe 2000. Conference Proceedings. S. 83–97, s.

den Aufbau des semantischen Netzes, das den zentralen innovativen Aspekt des Projektes darstellt. Dabei spiegeln sich 'alte' Überlegungen in einem neuartigen Konzept, an dessen Vielschichtigkeit und Komplexität gezeigt werden kann, welche Möglichkeiten es heute im Umgang mit verlegerischen Inhalten gibt.

2. Das Thomas-Mann-Projekt

Thomas Mann ist einer der wichtigsten Autoren des 20. Jahrhunderts. Sein umfangreiches Werk, in dem sich viele der großen Themen seiner Zeit widerspiegeln, gehört zum festen Kanon der deutschsprachigen Literatur. In diesem Jahr werden die ersten Bände der 'Großen Kommentierten Frankfurter Ausgabe der Werke Thomas Manns' (GKFA) erscheinen.⁵ Diese wissenschaftliche Studienausgabe wird die nicht kommentierten 'Gesammelten Werke Thomas Manns' (GW) als Referenzausgabe ablösen.⁶ Der Plan für diese neue umfangreiche Ausgabe war für den S. Fischer Verlag der Anlass, über eine neuartige Strategie im Umgang mit den Thomas-Mann-Texten im Verlag nachzudenken und schließlich das Thomas-Mann-Projekt zur Verwirklichung dieser Strategie auf den Weg zu bringen.

unter <http://www.gca.org/papers/xmleurope2000/papers/s09-02.html> (eingesehen am 10.3.2002); erneut publiziert in: Markup Languages: Theory & Practice 2/4 (2000), S. 353–365.; INGRID SCHMIDT/CAROLIN MÜLLER: Zaubernetz. Inhaltsstrukturen und Topic Maps als Potenzial neuer Informationstechnik. In: iX 11 (2000), S. 100–107, s. unter <http://www.heise.de/ix/artikel/2000/11/100> (eingesehen am 10.3.2002); INGRID SCHMIDT/CAROLIN MÜLLER: A world to discover. A topic map for Thomas Mann. In: XML Europe 2001. Conference Proceedings, s. unter <http://www.gca.org/papers/xmleurope2001/papers/html/s10-2.html> (eingesehen am 10.3.2002); INGRID SCHMIDT/CAROLIN MÜLLER: Das Thomas Mann Projekt. Eine literarische Edition auf neuen Wegen. In: Sprache in Hypermedia-Anwendungen. Hrsg. von HERMANN CÖLFEN, ANDREAS LIEBERT und ANGELIKA STORRER (OBST – Osnabrücker Beiträge zur Sprachtheorie 63 [2001]), S. 97–114. – Einführend zur Konzeption semantischer Netze vgl. STEVE PEPPER: Navigating Haystacks and discovering needles. Introduction in the new Topic Map Standard. In: Markup Languages 1/4 (1999), S. 47–74; HANS HOLGER RATH: Mozart oder Kugel. Mit Topic Maps intelligente Informationsnetze aufbauen. In: iX 12 (1999), S. 149–155.

⁵ In jährlichem Rhythmus werden bis 2015 insgesamt 58 Bände der Werke, Briefe und Tagebücher Thomas Manns (inkl. Kommentarbände) erscheinen; die Thomas-Mann-Texte werden von über 15 international renommierten Wissenschaftlern kommentiert.

⁶ Thomas Mann: Gesammelte Werke in 12 Bänden (Band 13: Nachträge). Frankfurt am Main 1960–1974.

2.1. Überblick über die Projektkonzeption

Die GKFA soll über Jahrzehnte hinweg einen einzigartigen Zugang zu den Werken Thomas Manns bieten. Bei einem auf eine so große Zeitspanne angelegten Projekt ist es in der heutigen Zeit selbstverständlich, dass die Werk- ausgabe aus einer Buchausgabe und einer elektronischen Ausgabe besteht. Beide Teile der Ausgabe sollen dabei hohen Qualitätsansprüchen genügen, langfristig nutzbar sein, den Möglichkeiten ihres Mediums gerecht werden und dabei Maßstäbe setzen. Bei einem Autor diesen Ranges – dazu seit Jahrzehnten einem der umsatzstärksten bei S.Fischer – gehört es genauso zur verlegerischen Aufgabe, die Konzeption und Aufbereitung der Inhalte für das elektronische Produkt zu betreuen, zu koordinieren und die damit verbundenen Funktionalitäts- und Präsentationsansprüche zu formulieren, wie dies auch beim gedruckten Buch selbstverständliche Tradition ist.

Diese hohen Anforderungen verlangen ein innovatives Konzept der Informationsorganisation: Es soll sich bis zum Abschluss der Ausgabe und darüber hinaus als tragfähig erweisen, keinen kurzfristigen technischen Entwicklungen unterliegen – sondern an internationalen Standards wie SGML/XML⁷ ausgerichtet sein – und trotzdem heute realisiert werden können. Das Konzept einer solchen Informationsarchitektur wird mit dem Aufbau eines Informationspools und eines semantischen Netzes umgesetzt.

Mit herkömmlichen Konzepten für elektronische Ausgaben literarischer Editionen ist dieses Konzept nicht zu verwirklichen. Sie bilden in der Regel die Buchausgabe elektronisch ab, ohne wesentlich neue Zugriffsmöglichkeiten außer der Volltextsuche zu bieten. Dabei wird die Volltextsuche den Ansprüchen des Benutzers meist nicht gerecht: Als Suchergebnis stehen häufig zu viele Fundstellen unsortiert nebeneinander; inhaltlich zentrale Stellen werden dabei nur dann gefunden, wenn auch das Suchwort im Text explizit vorkommt. Das hieße für eine Volltextsuche im Thomas-Mann-Werk, dass weder Schiller und Goethe in der 'Schweren Stunde' zu finden wären, noch wäre zu recherchieren, dass die 'Buddenbrooks' in Lübeck spielen, da die jeweiligen Suchwörter nicht im Text erwähnt sind.

Diese Beispiele zeigen, dass sich die Zugriffsmöglichkeiten der elektronischen Ausgabe nicht auf eine Volltextsuche beschränken, sondern die Textinhalte transparent machen sollte. Mit der Erschließung der Inhalte durch ein

⁷ SGML: Standard Generalized Markup Language, ISO-Standard 8879; XML: eXtensible Markup Language; eine gute Übersicht dazu bietet die Webseite von ROBIN COVER: The XML Cover Pages, s. unter <http://www.oasis-open.org/cover/sgml-xml.html> (eingesehen am 10.3.2002).

semantisches Netz kann nun gewährleistet werden, dass alle Vorkommen von 'Lübeck' zu finden sind, auch wenn das Wort 'Lübeck' nicht explizit im Text steht. Darüber hinaus können die Fundstellen nach Kategorien geordnet angezeigt werden, z. B.:

- fiktionaler Ort im belletristischen Werk,
- Bezugsort in Essays,
- Aufenthaltsort Thomas Manns,
- Absendeort im Brief.

Eine solche Form der Texterschließung ermöglicht außerdem neuartige Anfragen, wie statistische Auswertungen der von Thomas Mann am häufigsten zitierten Autoren, eine Liste der Zeitungen, in denen er am meisten publizierte u. v. m.

Die Informationsarchitektur muss konzipiert, aufgebaut und umgesetzt werden. Dabei ist es ein zentraler Punkt, der klaren und weitsichtigen Konzeption, der inhaltlichen Komplexität sowie dem zeitlich lang angelegten Rahmen der GKFA gerecht zu werden. Der Aufbau eines semantischen Netzes ermöglicht es, die Inhalte zu erschließen und für verschiedene Publikationsmedien zu öffnen. So kann eine elektronische Ausgabe entwickelt werden, die ein neuartiges Arbeitswerkzeug für Literaturwissenschaftler darstellt.⁸

2.2. Das semantische Netz

Der Aufbau eines semantischen Netzes war die größte Herausforderung im Thomas-Mann-Projekt. Es hat die Funktion einer Metastruktur und wird in SGML/XML-Notation als Topic Map exportiert.

Literarische Texte erweisen sich als besonders schwierig beim Aufbau eines semantischen Netzes. Enzyklopädien und Wörterbücher sind geprägt von festen Strukturen und Formulierungsvorgaben. In technischen Handbüchern wird in aller Regel zunächst eine Terminologie eingeführt, die anschließend konsistent verwendet wird. Literarische Texte verhalten sich ganz anders: Für sie ist die Sprache naturgemäß ein Mittel des künstlerischen Ausdrucks. Sie nutzen die Variationsmöglichkeit der Sprache, ihre Ambiguität und ihre Bildmöglichkeiten. So werden wie gesagt in der Erzählung 'Die schwere Stunde' die beiden Protagonisten Schiller und Goethe kein einziges

⁸ Siehe zu diesem Thema auch SÖREN STEDING: Warum noch drucken? Über die Nachteile gedruckter Editionen. In: Jahrbuch für Computerphilologie – online, s. unter <http://computerphilologie.uni-muenchen.de/jg01/steding.html> (eingesehen 10.3.2002).

Mal namentlich erwähnt. Statt dessen stehen die Phrasen 'der aus Jena' und 'der aus Weimar'. Gerade diese Variationsmöglichkeiten der Sprache machen ein schriftstellerisches Werk aus und erschweren zugleich seine automatisierte Inhaltserschließung.

Daher wurde schon im Vorfeld der Projektkonzeption klar, dass nicht die Thomas-Mann-Texte semantisch ausgezeichnet werden können, sondern die damit verknüpften Herausgeberkommentare. Dieser Entscheidung lag die Annahme zugrunde, dass diese wissenschaftlichen Texte auch in einer wissenschaftlichen Sprache formuliert sind. Dieser Sachverhalt war allerdings bei näherem Hinsehen nicht gegeben. Die Herausgebertexte sind zweifelsohne aufgrund ihres interpretatorischen Anliegens expliziter als die Texte Thomas Manns; daher erwähnen sie z.B. Goethe und Schiller auch namentlich. Sie sind allerdings in zu hohem Maße von einer poetischen Ausdrucksweise geprägt, um als formalisierte Texte angesehen werden zu können.⁹ Dies zeigt sich beispielsweise darin, dass Autor und Werktitel im Text meist nicht in formalisierter Form hintereinander stehen, sondern in eine Phrase wie 'der Autor X erwähnt in seinem jüngsten Werk über das Leben Thomas Manns ...' eingebunden sind. Dabei sind in den Kommentaren weder Werktitel noch Erscheinungsjahr genannt. Die genauen Angaben finden sich nur in der angehängten Bibliographie.

Auch in den Herausgebertexten ist daher eine semantische Textauszeichnung nur sehr begrenzt möglich. Stützt sich eine Recherche also nur auf die semantische Auszeichnung der Texte, sind ihre Möglichkeiten stark eingeschränkt. Werden jedoch sprachlich unterschiedlich dargestellte Sachverhalte in die formale Notation eines semantischen Netzes übertragen und dann mit den entsprechenden Textstellen verknüpft, kann gezielt – und darüber hinaus auf vielfältige Weise – auf die Inhalte zugegriffen werden.

Im Folgenden wird kurz die Konzeption des semantischen Netzes veranschaulicht.

⁹ Die von FRICKE durchgeführte Untersuchung von Wissenschaftssprache zeigt, dass literaturwissenschaftliche Texte überdurchschnittlich stark von der poetischen Sprache des Autors geprägt sind, auf dessen Werk sie sich beziehen. HARALD FRICKE: Die Sprache der Literaturwissenschaft. Textanalytische und philosophische Untersuchungen. München 1977.

2.2.1. Konzeption

Auf der formalen Ebene teilt sich das semantische Netz in (1) die Ebene der Begriffe und (2) die Ebene der Individuen. Auf der Begriffsebene werden die Begriffe und Relationen festgelegt und zu einem Begriffsnetz verknüpft. Hier wird auch bestimmt, welche der Begriffe individuenfähig sind und welche Relationen gelten. Dabei ist es von zentraler Bedeutung, welche Basisbegriffe und Basisrelationen festgelegt werden, da sie die Möglichkeiten einer späteren Ausdifferenzierung festlegen. Eben diese graduellen Erweiterungsmöglichkeiten stellen das Zukunftspotenzial des semantischen Netzes dar. Auf diese Weise entsteht ein abstraktes Modell von Begriffen und Relationen, das anschließend als Raster für die Einordnung und Verknüpfung der Individuen dient.

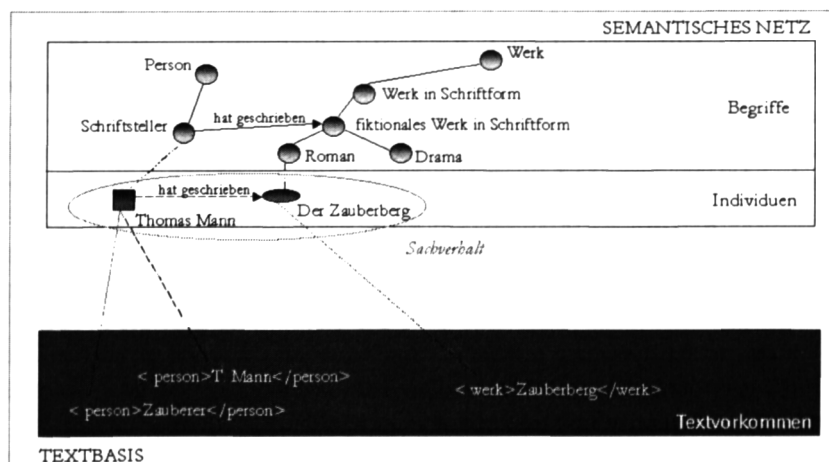


Abb. 1: Veranschaulichung des semantischen Netzes

Zwischen Individuen und Begriffen gilt die Instanz-Relation; sie ist eine 'ist'-Relation. In diesem Sinne kann man in der Veranschaulichung des semantischen Netzes lesen: Thomas Mann ist ein Schriftsteller. Um einen solchen Sachverhalt auf der Individuenebene ausdrücken zu können, muss auf der Begriffsebene die Möglichkeit angelegt sein, Personen nach ihren verschiedenen Rollen und Funktionen, als Familienmitglieder, Verleger, Schriftsteller, Politiker u.a. einzuordnen. Erst dann kann das Individuum 'Thomas Mann' mit dem Begriff 'Schriftsteller' eine Instanz-Relation bilden. Entsprechendes gilt für 'Der Zauberberg ist ein Roman' und 'Ein Roman ist ein Werk in Schriftform'. Das Individuum 'Der Zauberberg' wird so mit dem Begriff 'Roman' verbunden. Erst wenn diese beiden Individuen mit ihren

Begriffen verbunden sind, können sie auf der Individuenebene mit der 'hat-geschrieben'-Relation verbunden werden. Mit einer solchen Verknüpfung wird ein sogenannter Sachverhalt festgehalten; in diesem Fall 'Thomas Mann hat den Zauberberg geschrieben'. Die 'hat-geschrieben'-Relation gilt auf der Individuenebene nur dann für alle Schriftsteller und ihre Werke, wenn dies auf der Begriffsebene vorgesehen ist. Auf der Begriffsebene ist die 'hat-geschrieben'-Relation mit ihren Eigenschaften definiert. Beispielsweise wird bestimmt, dass sie für das Begriffspaar 'Schriftsteller' und 'Werk in Schriftform' gilt und dass sie vererbbar ist, d.h. auch für alle Unterbegriffe dieses Begriffspaares gilt. Damit gilt die 'hat-geschrieben'-Relation für alle Individuen, die mit diesen Begriffen bzw. ihren Unterbegriffen verknüpft sind.

Zusätzlich zu der Verknüpfung der Individuen mit den Begriffen werden die Individuen mit Metadaten angereichert; über den Aufbau der Begriffsebene wird ein Ordnungssystem für ein Wissensgebiet festgelegt; das Sachwissen wird über die Verknüpfung der Individuen zu Sachverhalten abgebildet. So bildet das semantische Netz auf unterschiedlichen Ebenen unterschiedliche Bereiche eines Wissensgebietes ab.

Dieses führt – sowohl für den Redakteur als auch für den Nutzer der elektronischen Version – zu bislang nicht gekannten Suchmöglichkeiten. Bezogen auf das oben gezeigte Beispiel kann er allgemein nach Schriftstellern suchen oder auch eingeschränkt danach, ob sie z.B. Romane oder Lyrik geschrieben haben. Er kann sich aber auch sämtliche im semantischen Netz vorhandenen Personen auflisten lassen. Geht man über die im Beispiel gezeigte Modellierung hinaus, so wird es auch möglich sein, nach allen Personen zu suchen, die vor 1900 geboren sind, oder nach allen Familienangehörigen von Thomas Mann.

Wie in Abbildung 1 veranschaulicht wurde, werden die Individuen schließlich mit der Textbasis verknüpft. Dabei ist das Individualobjekt 'Thomas Mann' mit allen Textstellen verbunden, die Thomas Mann nennen oder meinen. Es ist also unerheblich, ob im Text nun die Zeichensequenz 'Thomas Mann', 'T. Mann' steht oder ob vom 'Zauberer' gesprochen wird, wie Thomas Mann von seinen Kindern genannt wurde.

Bei der Entwicklung des semantischen Netzes sind daher zwei Aufgabenbereiche zu unterscheiden: die Modellierung des Begriffsnetzes und der konkrete Aufbau der Individuenebene; auf Letzteres wird im Folgenden näher eingegangen.

2.2.2. *Aufbau der Individuenebene*

Neben der Entwicklung des Begriffsnetzes ging es darum, die Individuenebene als Individuennetz aufzubauen. Dabei müssen die Individuen sowohl mit dem Begriffsnetz als auch untereinander verknüpft werden. Für den Aufbau gibt es verschiedene Möglichkeiten. Die Individuenebene ausschließlich manuell mit Individualobjekten zu füllen und diese zu Sachverhalten zu verknüpfen, ist sehr arbeitsintensiv und damit zeit- und kostenaufwändig. Weit- aus sinnvoller ist es, in einem ersten Schritt einen Grundbestand an Individuen zu importieren, sie automatisch mit dem Begriffsnetz und soweit wie möglich zu Sachverhalten zu verknüpfen und das Ergebnis anschließend redaktionell nachzubearbeiten. Für eine solche Vorgehensweise beim Aufbau der Individuenebene können die bereits bestehenden Register zum Werk Thomas Manns herangezogen werden. Weitere Quellen für den Aufbau der Individuenebene sind die in der Redaktion erstellten Metadaten zu den einzelnen Textsorten, die Herausgeberkommentare zu den Thomas-Mann-Texten, die den Kern der Editionsarbeit bilden, sowie andere Quellen, die nicht in direktem Zusammenhang mit dem Werk Thomas Manns stehen.

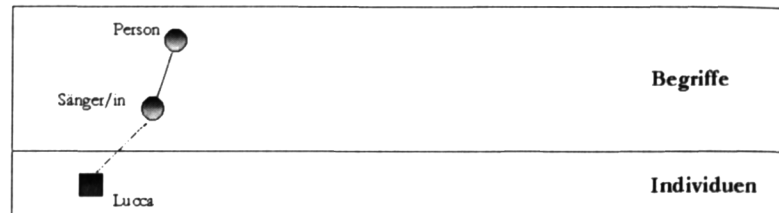
2.2.2.1. *Register*

Register für den Aufbau und die Verknüpfung des Individuennetzes auszunutzen, bietet sich aus zweierlei Gründen an:

- Inhaltlich: Register sind, im Gegensatz zu Volltextindizes, von kundigen Menschen erstellt. Sie nehmen nur die für den Text relevanten Begriffe auf, über die für den Leser eine festgelegte Zugriffsstruktur auf die Inhalte entsteht. Ein Register in einem Buch stellt immer eine Auswahl aus vielen möglichen Registern dar, die aus Platzgründen getroffen werden muss. Semantische Netze stellen quasi ‘fortgeschrittene’ Register dar, die einen flexiblen, auf die jeweilige Benutzungssituation zugeschnittenen Zugang ermöglichen.
- Formal: Register sind stark formalisiert und daher gut automatisch auszuwerten. Sie bestehen in der Regel nicht nur aus Registerbegriffen, sondern lassen über ihre formale Struktur auch Sachverhalte erkennen.

Der formale Aspekt der Auswertbarkeit der Registereinträge soll an zwei Beispielen veranschaulicht werden.

Lucca (Sängerin)



Lot (biblische Gestalt)

Lusitania (Schiff)

Adonis (Gestalt der griechischen Mythologie)

Abb. 2: Instanz-Relation

Die in Abbildung 2 gezeigten Registereinträge nennen jeweils einen Individualnamen, der über den dahinter stehenden Ausdruck in Klammern näher bestimmt wird. In allen diesen Fällen gilt die Instanz-Relation für den Individualnamen und den Text in der Klammer: 'Lucca ist eine Sängerin', 'Lot ist eine biblische Gestalt', 'Lusitania ist ein Schiff' und 'Adonis ist eine Gestalt der griechischen Mythologie'.

Das zweite, in Abbildung 3 aufgeführte Beispiel geht noch einen Schritt weiter. In den hier gezeigten Registereinträgen kann nicht nur die Instanz-Relation ausgewertet werden, sondern es können auch Sachverhalte abgeleitet werden.

Der erste Eintrag besteht aus einer recte gesetzten Zeile und zwei eingezogenen kursiven Zeilen. Dieses formale Muster bedeutet, dass es sich um eine Person und um das von ihr geschaffene Werk handelt. Den Personennamen kann man aufgrund des Kommas in Vor- und Familiennamen gliedern. Aus der formalen Struktur des Eintrags kann man allerdings weder die Rolle der Person, noch die Art des Werks ableiten. Daher kann man auch nichts Näheres über die Art des Schaffensprozesses sagen. Zusammengefasst lassen sich diesem Registereintrag folgende Informationen entnehmen:

- Gotthold Ephraim Lessing ist eine Person,
- Lessing ist ein Familienname,
- Gotthold ist ein Vorname,

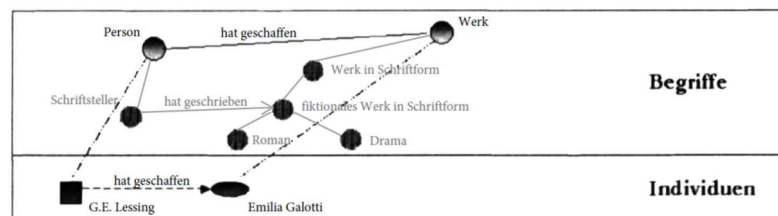
- Ephraim ist ein Vorname,
- Emilia Galotti ist ein Werk,
- Nathan der Weise ist ein Werk,
- G. E. Lessing hat Emilia Galotti geschaffen,
- G. E. Lessing hat Nathan der Weise geschaffen.

Genauere Zuordnungen wie 'G. E. Lessing ist ein Schriftsteller' oder 'Emilia Galotti ist ein Drama' und damit 'Lessing hat Emilia Galotti geschrieben' müssen manuell vorgenommen werden. Für den Eintrag zu Richard Wagner gilt Entsprechendes. Hier sind über die doppelte Einrückung darüber hinaus noch die einzelnen Werkteile des 'Ring des Nibelungen' abzuleiten.

Lessing, Gotthold Ephraim

Emilia Galotti

Nathan der Weise



Wagner, Richard

Der Ring des Nibelungen

Das Rheingold

Die Walküre

Siegfried

Götterdämmerung

Abb. 3: Sachverhalte

Zum Werk Thomas Manns liegen mehrere Register vor, einige davon in elektronischer Form. Für den ersten Schritt wurde das 'Verzeichnis der erwähnten Namen, Personen und fremden Werke' aus der von HERMANN KURZKE und STEPHAN STACHORSKI herausgegebenen Thomas-Mann-Essay-Ausgabe ausgewählt,¹⁰ aus der auch die eben gezeigten Beispiele entnommen sind. Für

¹⁰ Dieses Register ist abgedruckt in: Thomas Mann: Essays. Hrsg. von HERMANN KURZKE und STEPHAN STACHORSKI. Band 6: Meine Zeit, 1945–1955. Frankfurt a.M. 1993, S. 626–720.

dieses Register sprach besonders, dass es sehr konsistent ausgezeichnet ist und in RTF-Format vorliegt. An ihm wird nun beispielhaft gezeigt, wie Register für das semantische Netz aufbereitet werden können. Diese Aufarbeitung erfolgt in drei Schritten, die nachfolgend näher beschrieben werden:

1. Konvertierung nach XML,
2. Erstellung von Teilregistern,
3. Bearbeitung der Teilregister.

Erster Schritt: Konvertierung nach XML

Zunächst wird das Register auf seine Registerbegriffe reduziert, d.h. Bandangaben, Seitenzahlen u.ä. werden entfernt, da sie irrelevant für das semantische Netz sind. Nachfolgend wird eine exemplarische Zusammenstellung von Registereinträgen gezeigt:

Ady, Endre (Andreas)
 Äsop (Aisopos)
 Benesch (Benes), Eduard
 Casanova (Giacomo Girolamo Casanova de Seingalt)
 Conti (Maler), s. Gotthold Ephraim Lessing, 'Minna von Barnhelm'
 Correggio (eigtl. Antonio Allegri)
 Goncourt, Brüder (Edmond-Louis-Antoine Huot de G. und Jules Alfred Huot de G.)
 Jupiter (römischer Gott)
 Jupiter (Planet)
 Simon & Schuster, Verlag (New York)

Im ersten Analyseschritt wird die Gestaltung der Registereinträge ausgewertet, indem das Zusammenspiel von Layout und Typographie im gedruckten Register beschrieben wird. Aus den sich daraus ergebenden formalen Mustern werden inhaltliche Zusammenhänge und ihre Benennungen abgeleitet. Das Analyseergebnis bildet die Grundlage für die regelbasierte automatische Aufbereitung des Registers. Das Ergebnis dieser Konvertierung ist ein semantisch ausgerichtetes wohlgeformtes XML-Dokument. Warum XML ausgewählt wurde und warum dabei wohlgeformtes XML validem XML vorgezogen wurde, hat drei Gründe:

1. Bearbeitungsformat: XML wird als Übergangsformat von den Satzdaten hin zum semantischen Netz eingesetzt. Dabei ist es vorteilhaft, dass mit XML die Semantik der Registereinträge transparent gemacht werden kann, so dass sie gut manuell bearbeitet werden können. Das Format ist dar-

über hinaus geeignet, das Register nach der Bearbeitung automatisch in die Struktur des semantischen Netzes zu überführen: Die XML-Auszeichnung nimmt dabei Bezug auf die Begriffsebene, der Elementinhalt entspricht dem Individualobjekt; aus der Reihenfolge der Auszeichnungen lassen sich Sachverhalte ableiten.

2. Flexibilität: Da ein Register im Vorfeld in aller Regel nicht in umfassender Weise analysiert werden kann, ist es nicht vollständig absehbar, ob und wenn ja welche Elemente bei der manuellen Bearbeitung der Register zusätzlich benötigt werden. So stellte sich beispielsweise bei der manuellen Nachbearbeitung heraus, dass auch ein Element PLANET gebraucht wird, da Jupiter als römischer Gott und als Planet im Register verzeichnet ist. Dieses Element kann bei wohlgeformtem XML einfach angelegt, gegebenenfalls mit Attributen ergänzt und gleich verwendet werden. Dadurch kann das Register äußerst flexibel bearbeitet werden.

3. Kostenersparnis: Eine detaillierte Analyse des Registers und eine darauf basierende DTD-Entwicklung ist zeit- und damit auch kostenintensiv. Die DTD wird in aller Regel von einer anderen Person erstellt und bearbeitet als die Register. Daher fallen erhöhte Kosten bei der Registerbearbeitung an, wenn die DTD um zusätzliche Elemente erweitert werden muss: für die DTD-Pflege und für die Wartezeit. Denn es muss bei validem XML in Kauf genommen werden, dass das Register erst weiter bearbeitet werden kann, wenn die neue DTD-Version erstellt ist. Weder die Strenge des Formats noch die Vorgehensweise würde dem Wunsch nach einer flexiblen Arbeitsumgebung entsprechen.

Zweiter Schritt: Erstellung von Teilregistern

Das als Konvertierungsergebnis entstandene wohlgeformte XML-Register stellt die Ausgangsbasis für die Erstellung von Teilregistern dar. Diese Teilregister werden mit einem eigens dafür entwickelten SGML/XML-Scanner erstellt. Mit diesem Programm kann die formale Struktur beliebiger SGML- bzw. XML-Dateien analysiert werden und es können gezielt Einträge nach bestimmten inhaltlichen Kriterien herausgelöst werden. Das Gesamtregister verfügt über mehr als 1000 Einträge, die alphabetisch, aber nicht hinsichtlich ihrer inhaltlichen Belange sortiert sind. Durch den bei der Erstellung von Teilregistern eingesetzten Filtermechanismus entstehen ganz spezifische Sichten auf das Gesamtregister, die eine Überarbeitung der Konvertierung vereinfachen. So gruppiert ein mögliches Teilregister beispielsweise alle Personen mit ihren Verwandtschaftsbeziehungen.

Dritter Schritt: Bearbeitung der Teilregister

Die Bearbeitung der Registereinträge wird anhand des Teilregisters aller Personen ohne Werke exemplarisch dargestellt. Erstellt man dieses Teilregister, so wird schnell deutlich, dass bei der Konvertierung Personen nicht von Institutionen unterschieden werden können. Sie sind im RTF-Format typographisch gleich behandelt und verfügen über keine zusätzlichen Charakteristika, die es ermöglichen, sie automatisch auseinanderzuhalten. Da in dem hier zugrunde gelegten Register wesentlich mehr Personen als Institutionen vorkommen, werden bei der Konvertierung alle Institutionen unter Personen subsummiert; diese automatisch vorgenommene Auszeichnung muss anschließend manuell korrigiert werden.

```
<ENTRY>
<PERSON>Bayerische Akademie der Sch&ouml;n K&uuml;nste</PERSON>
</ENTRY>
```

wird korrigiert zu:

```
<ENTRY>
<INSTITUTION>Bayerische Akademie der Sch&ouml;n K&uuml;nste
</INSTITUTION>
</ENTRY>
```

Diese Auszeichnungsstrategie soll nun beispielhaft an einigen Registereinträgen verdeutlicht werden. Jedes Beispiel wird eingeleitet durch die Darstellung des gedruckten Registereintrags. Nachfolgend wird die automatisch konvertierte XML-Struktur gezeigt und anschließend die Korrekturen, die durch Erläuterungen ergänzt sind. Die Beispiele stehen darüber hinaus exemplarisch für die komplexen Anforderungen, die an die Auszeichnung der Registereinträge gestellt werden, damit sie später automatisch in das semantische Netz eingelesen werden können.

Ady, Endre (Andreas)

```
<ENTRY><PERSON ALTNAMES='Andreas'>Ady, Endre</PERSON></ENTRY>
<ENTRY><PERSON ALTNAMES-VOR='Andreas'>Ady, Endre</PERSON></ENTRY>
```

Dieser Eintrag kann durch eine einfache Umbenennung des Attributs ALTNAMES zu ALTNAMES-VOR korrigiert werden. Damit wird verdeutlicht, dass es sich bei 'Andreas' um einen alternativen Vornamen handelt.

Benesch (Benes), Eduard

```
<ENTRY><PERSON ALTNAMES='Benes'>Benesch, Eduard</PERSON></ENTRY>
<ENTRY><PERSON ALTNAMES-NACH='Benes'>Benesch, Eduard</PERSON></ENTRY>
```

Bei diesem Eintrag wird wie bei dem vorhergehenden verfahren, nur dass es sich in diesem Fall um einen alternativen Familiennamen handelt.

Jupiter (Planet)

```
<ENTRY><PERSON ALTNAME='Planet'>Jupiter</PERSON></ENTRY>
<ENTRY><PLANET>Jupiter</PLANET></ENTRY>
```

Aufgrund der formalen Struktur wird 'Planet' bei diesem Eintrag als Namensvariante von Jupiter erkannt, der seinerseits als PERSON eingeordnet wird. Hier wird eine grundsätzliche Änderung vorgenommen, indem 'Jupiter' mit dem Element PLANET ausgezeichnet wird.

Jupiter (römischer Gott)

```
<ENTRY><MYTHG TYPE='römischer Gott'>Jupiter</MYTHG></ENTRY>
```

Obwohl die rein formale Struktur hier identisch ist mit der des vorherigen Registereintrags, wird 'Jupiter' hier korrekt als mythologische Gestalt erkannt. Dies ist möglich, weil in der Klammer ein Mehrwortausdruck mit 'Gott' als Schlüsselwort steht. Bei der Konvertierung kann aufgrund des Mehrwortausdrucks ausgeschlossen werden, dass es sich um eine mythologische Gestalt handelt; im Christentum als monotheistischer Religion wird 'Gott' nicht durch einen Zusatz näher bestimmt.

Äsop (Aisopos)

```
<ENTRY><PERSON ALTNAME='Aisopos'>&Auml;sop</PERSON></ENTRY>
<ENTRY><PERSON>&Auml;sop</PERSON><ALTNAME>Aisopos</ALTNAME></ENTRY>
```

Bei 'Aisopos' handelt es sich um eine Namensalternative, die den Personennamen 'Äsop' vollständig ersetzt. Sie wird deshalb nicht ins Attribut zu PERSON gestellt, sondern von dem Element ALTNAME umklammert.

Correggio (eigtl. Antonio Allegri)

```
<ENTRY><PERSON>Correggio</PERSON>
<ALTNAME TYPE='eigtl.'>Antonio Allegri</ALTNAME> </ENTRY>
```

Dieser Eintrag wird korrekt konvertiert, da die Namensvariante in der runden Klammer durch das Schlüsselwort 'eigtl.' eingeleitet wird; es wird als Wert des Attributs TYPE mitgeführt. Alle vollständigen Namensvarianten werden im gedruckten Register mit 'eigtl.' oder 'Pseudonym:' eingeleitet. Das letzte Wort des Elementinhalts von ALTNAME ist der Familienname, so dass der Name ggf. automatisch invertiert werden kann.

Casanova (Giacomo Girolamo Casanova de Seingalt)

```
<ENTRY><PERSON ALTNAME='Giacomo Girolamo Casanova de Seingalt'>Casano-
va</PERSON></ENTRY>
<ENTRY><PERSON>Casanova</PERSON>
<ALTNAME>Giacomo Girolamo Casanova de Seingalt</ALTNAME></ENTRY>
<ENTRY><PERSON>Casanova</PERSON>
<ALTNAME>Casanova de Seingalt, Giacomo Girolamo</ALTNAME></ENTRY>
```

Auch beim Registereintrag zu 'Casanova' handelt es sich in der runden Klammer um eine vollständige Namensvariante, weswegen wie im vorheri-

gen Beispiel verfahren wird. Im Gegensatz dazu lassen sich aber hier Vornamen und Familienname unterscheiden. Allerdings ist der Familienname nicht einfach das letzte Wort, sondern umfasst die letzten drei Wörter. Er wird daher in einem zweiten Schritt manuell in Vor- und Nachnamen aufgeteilt.

Goncourt, Brüder (Edmond-Louis-Antoine Huot de G. und Jules Alfred Huot de G.)

```
<ENTRY>
<PERSON ALTNAME='Edmond-Louis-Antoine Huot de G. und Jules-Alfred Huot de G.' >Goncourt, Br&uuml;der</PERSON></ENTRY>
<ENTRY><PERSONENGRUPPE ALTNAME='Edmond-Louis-Antoine Huot de G. und Jules-Alfred Huot de G.' >Goncourt, Br&uuml;der</PERSONENGRUPPE></ENTRY>
<ENTRY><PERSONENGRUPPE>Goncourt, Br&uuml;der</PERSONENGRUPPE>
<PERSON>Edmond-Louis-Antoine Huot de G.</PERSON>
<PERSON>Jules-Alfred Huot de G.</PERSON></ENTRY>
<ENTRY><PERSONENGRUPPE>Goncourt, Br&uuml;der</PERSONENGRUPPE>
<PERSON>Edmond-Louis-Antoine Huot de Goncourt</PERSON>
<PERSON>Jules-Alfred Huot de Goncourt</PERSON></ENTRY>
```

Die Korrektur dieses Registereintrags ist ungleich aufwändiger als die der bisher gezeigten. Die automatische Konvertierung erkennt eine Person mit alternativem Namen; die Struktur entspricht vollständig der des Beispiels 'Ady, Endre (Andreas)'. Inhaltlich handelt es sich aber um ein Brüderpaar und nicht um eine Einzelperson. Deshalb wurde als erstes das Element PERSONENGRUPPE neu definiert und anstatt der Auszeichnung PERSON gesetzt. Bei dem Attributwert von ALTNAME handelt es sich weder um einen alternativen Vor- bzw. Familiennamen noch um eine vollständige Namensvariante. Vielmehr stehen hier die Personen, aus denen sich die Personengruppe zusammensetzt; sie werden im nächsten Schritt als einzelne Personen gefasst. Auffällig ist, dass der Familienname im Elementinhalt von PERSON als 'G.' abgekürzt ist. Intellektuell kann diese Abkürzung sehr einfach als 'Goncourt' aufgelöst werden, da der Familienname Teil des Elementinhalts von PERSONENGRUPPE ist; automatisch ist dies jedoch nicht ohne weiteres möglich. Damit eine sinnvolle automatische Invertierung der Personennamen möglich wird, muss der Familienname in beiden Fällen manuell ergänzt werden.

Simon & Schuster, Verlag (New York)

```
<ENTRY>
<PERSON ALTNAME='New York' >Simon & Schuster, Verlag </PERSON></ENTRY>
<ENTRY><ORGANISATION ALTNAME='New York' >Simon & Schuster, Verlag
</ORGANISATION></ENTRY>
<ENTRY><ORGANISATION ORT='New York' >Simon & Schuster, Verlag
</ORGANISATION></ENTRY>
<ENTRY>
```

```
<ORGANISATION TYP='Verlag' ORT='New York'>Simon & Schuster
</ORGANISATION></ENTRY>
```

Hier wird aufgrund des Layouts und der Formatierung bei der Konvertierung eine Organisation als Person klassifiziert und der Standort der Organisation als Namensvariante. Daher wird zunächst der Elementname korrigiert und anschließend wird ein neues Attribut ORT eingeführt; so bleibt der Standort eng mit der Organisation verbunden. Darüber hinaus ist der Organisationstyp an den Namen der Organisation angehängt. Er wird in den Attributwert des TYP-Attributs gestellt. Damit verfügt das Element ORGANISATION über zwei Attribute, die die Organisation näher bestimmen.

Nachdem das Register vollständig überarbeitet ist, werden die Registereinträge über einen automatisierten Prozess in das semantische Netz eingelesen; aus den semantischen Auszeichnungen der Einträge werden dabei so viele Sachverhalte wie möglich abgeleitet. Dabei sind im gedruckten Register nicht alle Sachverhalte so differenziert typographisch und layouttechnisch unterschieden, wie dies im semantischen Netz notwendig ist (vgl. Abbildung 3 und die Erläuterungen dazu). Die nach der Konvertierung und dem Import nur unspezifisch umgesetzten Sachverhalte müssen dann von der Redaktion manuell im semantischen Netz spezifischer zugeordnet werden. Eine weitere redaktionelle Aufgabe ist die ebenfalls manuelle Vervollständigung von Sachverhalten im semantischen Netz, die nicht über die Registereinträge abgebildet waren. Der Import der Registereinträge und die manuelle Ergänzung wird anhand eines Beispiels veranschaulicht:

```
Conti (Maler), s. Gotthold Ephraim Lessing, 'Minna von Barnhelm'
<ENTRY><FIGUR ROLLE='Maler'>Conti</FIGUR>
<LINK TYPE='s.'><PERSON>Lessing, Gotthold Ephraim</PERSON>
<WERK><TITEL>Minna von Barnhelm</TITEL></WERK></LINK></ENTRY>
```

Der gedruckte Registereintrag liegt nach der manuellen Korrektur der automatischen Konvertierung in XML vor. Diese semantischen XML-Auszeichnungen werden in die Struktur des semantischen Netzes überführt und können dann im Netzeditor von K-Infinity¹¹ angezeigt (Abbildung 4) und anschließend manuell erweitert (Abbildung 5) werden.

¹¹ K-Infinity ist eine Software der Firma intelligent views GmbH zum Aufbau semantischer Netze.

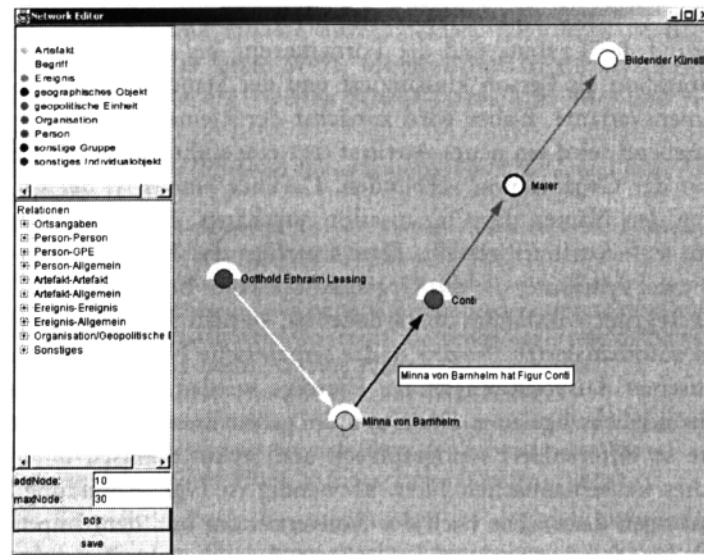


Abb. 4: Importierter Registereintrag

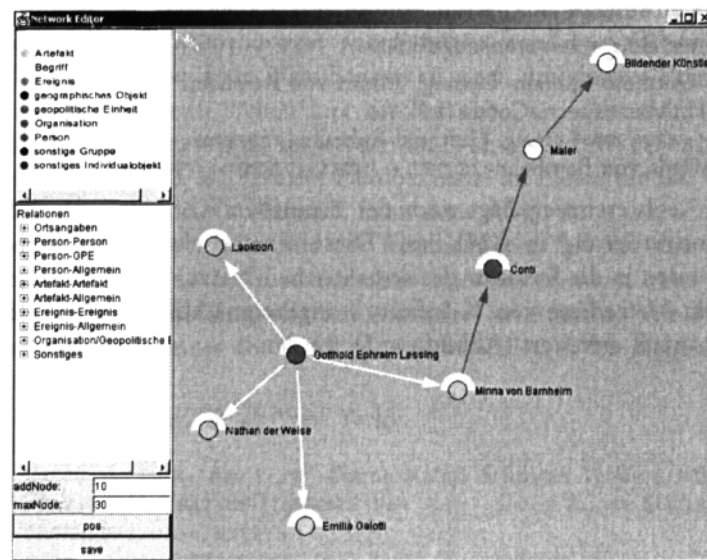


Abb. 5: Vervollständigung des importierten Registereintrags

2.2.2.2. *Metadaten*

Allen GKFA-Texten und -Dokumenten werden bei der redaktionellen Bearbeitung Metadaten beigegeben. Für die GKFA werden drei verschiedene Typen von Metadaten unterschieden:

1. Bibliographische Angaben:
 - allgemeine bibliographische Angaben,
 - textsortenspezifische bibliographische Angaben,
 - ausgabenspezifische bibliographische Angaben.
2. Textsortenspezifische Kategorisierungen.
3. Verschlagwortungen.

Jedes Dokument im Informationspool verfügt über Metadaten in SGML/XML: Es gibt Metadaten zu Romanen, Essays, Briefen, zu den Herausgeberkommentaren und zu den Archivmaterialien. Sie können über automatisierte Importroutinen ganz oder teilweise in das semantische Netz eingelesen werden. Damit werden auf der Individuenebene Sachverhalte zu den einzelnen Dokumenten abgebildet, wie beispielsweise die Entstehungszeit eines Werks, der Verlag des Erstdrucks, die nähere Spezifikation des Essaytyps oder auch um welchen Medientyp es sich bei Archivmaterialien handelt. Die Bandbreite reicht von einfachen bibliographischen Angaben bis hin zu detaillierten textsortenspezifischen Kategorisierungen.

So gibt es zu Romanen nur wenige Metadaten; die Metadaten zu Briefen und Essays sind hingegen wesentlich umfassender und komplexer. Bei den Briefen werden die zugehörigen Regesten herangezogen;¹² Briefe haben daher die ausführlichsten Metadaten.

2.2.2.3. *Herausgeberkommentare*

Während die Register und die Metadaten so aufbereitet werden können, dass sie automatisch in die Struktur des semantischen Netzes zu überführen sind, ist ein solches Vorgehen für die Herausgeberkommentare nicht denkbar. Dabei stellen sie die inhaltlich wichtigste Quelle für den Aufbau der Individuenebene dar. Ihre Auswertung ist jedoch sehr zeit- und kostenintensiv.

¹² Die Regesten zu den Thomas-Mann-Briefen wurden zwischen 1976 und 1987 von HANS BÜRGIN und HANS-OTTO MAYER bearbeitet und herausgegeben. Die Briefe Thomas Manns. Regesten und Register. Hrsg. von HANS BÜRGIN und HANS-OTTO MAYER, unter Mitarbeit von YVONNE SCHMIDLIN (Thomas-Mann-Archiv Zürich). Frankfurt 1976–1987.

So muss von sachverständigen Bearbeitern eine manuelle Bearbeitung der Texte vorgenommen werden: Die Texte müssen gelesen, verstanden und ausgezeichnet werden. Fordert das Lesen und Verstehen inhaltliche und literaturwissenschaftliche Kenntnisse, so ist für die Textauszeichnung ein Grundverständnis für die formale und verlagsstrategische Projektkonzeption erforderlich und soviel technisches Verständnis, um mit der K-Infinity-Software umgehen zu können. Nur dann, wenn dem Bearbeiter klar ist, warum welche Elemente zur Verfügung stehen und warum sie in der besonderen Granularität modelliert sind, kann er die Texte konsistent und sinnvoll auszeichnen. Und nur dann können sie beim Import bestmöglich für den Aufbau der Individuenebene verwendet und für die Verknüpfung der Individuenebene mit der Textbasis ausgenutzt werden.

Eine der Hauptschwierigkeiten bei den Herausgeber-texten ist ihre Sprache: Sie ist wenig formalisiert, wodurch häufig diskontinuierliche Wortfolgen, Ellipsen u. ä. auftauchen, die eigentlich von einem Element umklammert werden sollten, im Text aber nicht umklammert werden können. Daher könnte eine andere Strategie im Umgang mit den Herausgeber-texten sein, die Texte nicht auszuzeichnen, sondern die ihnen zu entnehmenden Informationen direkt im semantischen Netz als Sachverhalte abzubilden und diese dann mit der entsprechenden Textstelle zu verknüpfen.

Anhand des Herausgeberkommentars zu einem Thomas-Mann-Essay sollen die angesprochenen Schwierigkeiten deutlich gemacht werden. Das Kommentarbeispiel wurde hinsichtlich der in ihm enthaltenen Individuen und Sachverhalte analysiert; die Individuen wurden zu Sachverhalten ergänzt. Alle Sachverhalte sind im Anschluss an den Herausgeberkommentar aufgelistet.

Essay-Kommentar

Ilse Martens, die Schwester von Thomas Manns als 'Hans Hansen' in die Literaturgeschichte eingegangenen Jugendfreund Armin Martens (und Vorbild für das 'Fräulein Isenschubbe' in 'Königliche Hoheit'), besaß ein Album, das entsprechend einer zeittypischen Mode unter der Überschrift "Erkenne Dich selbst!" eine Reihe vorgedruckter Fragebögen enthielt. Unter den Besuchern, die dort ihre Antworten eintrugen, waren außer Thomas Mann sein Freund Paul Ehrenberg, John Eckhoff (der als 'Johnny Bishop' in der Erzählung 'Wie Jappe und Do Escobar sich prügelten' wiederauftaucht), Graf Vitzthum, der Mitautor des 'Frühlingssturm', und die Schwester Julia Mann. Auch wenn gelegentlich ein etwas renommistischer Ton bemerkbar ist, geben Thomas Manns Antworten doch ein erstaunlich plastisches und wahrheitsgemäßes – jedenfalls durch andere Zeugnisse derselben Zeit beglaubigtes – Selbstporträt.

Sachverhalte

Person (Thomas Mann) schafft Essay (Erkenne Dich Selbst)
 Person (Thomas Mann) schafft Roman (Königliche Hoheit)
 Person (Thomas Mann) schafft Roman (Tonio Kröger)
 Person (Thomas Mann) schafft Erzählung (Wie Jappe und Do Escobar sich prügelten)

Person (Thomas Mann) ist bekannt mit Person (Ilse Martens)
 Person (Thomas Mann) ist bekannt mit Person (Armin Martens)
 Person (Thomas Mann) ist bekannt mit Person (Paul Ehrenberg)
 Person (Thomas Mann) ist bekannt mit Person (Graf Vitzthum)
 Person (Thomas Mann) ist bekannt mit Person (John Eckhoff)

Person (Ilse Martens) ist Geschwister von Person (Armin Martens)
 Person (Thomas Mann) ist Geschwister von Person (Julia Mann)

Person (Ilse Martens) ist Vorlage für Figur (Fräulein Isenschibbe)
 Figur (Fräulein Isenschibbe) ist Figur in Roman (Königliche Hoheit)
 Person (Armin Martens) ist Vorlage für Figur (Hans Hansen)
 Figur (Hans Hansen) ist Figur in Roman (Tonio Kröger)
 Person (John Eckhoff) ist Vorlage für Figur (Johnny Bishop)
 Figur (Johnny Bishop) ist Figur in Erzählung (Wie Jappe und Do Escobar sich prügelten)

Person (Graf Vitzthum) steht in Beziehung mit Artefakt (Frühlingssturm)

Wie aus dem Beispiel zu ersehen ist, können aus einem relativ kurzen Kommentartext eine ganze Reihe von Sachverhalten abgeleitet werden. Dies kann jedoch nur intellektuell geschehen; ein Automatismus ist hier nur schwer vorstellbar.

2.2.2.4. Weitere Quellen

Als vierte Gruppe können für den Aufbau der Individuenebene Quellen genutzt werden, die nicht in unmittelbarer Verbindung mit dem Werk Thomas Manns stehen. Über sie können für das Wissensgebiet relevante Informationen vervollständigt werden, über die die GKFA nur punktuell Auskunft gibt. Diese Arbeiten sind von der Datenbasis der GKFA abgelöst und können daher als separate Arbeitspakete definiert und an externe Dienstleister vergeben werden. So können beispielsweise aus 'Kindlers Literaturlexikon' die Lebensdaten und die Werke von Schriftstellern extrahiert werden.¹³ Werden diese Arbeiten auf Basis einer einfachen wohlge-

¹³ Geht man bei der Datenextraktion von der Druckausgabe [Kindlers neues Literaturlexikon in 21 Bänden. Hrsg. von WALTER JENS. München 1988–1998] aus, so ist damit ein vollständig manueller Prozess verbunden; legt man die CD-ROM-Version

formten oder auch validen XML-Struktur gemacht, dann können die Arbeitsergebnisse über automatisierte Routinen in das semantische Netz eingelesen werden.

3. Ausblick

Ein Verlag muss sich heute von der an das Buchparadigma gebundenen Herangehensweise an die Produktion literarischer Ausgaben ablösen, wenn er die Weichen für die Zukunft stellen will. Im Zentrum stehen dann die Inhalte; das Buch macht Platz, rückt aus der Mitte in Richtung Rand und gesellt sich zu den anderen Publikationsmedien. Die Kernkompetenz des Verlags, die Buchpublikation, bewegt sich damit ebenfalls aus der Mitte weg. Dabei muss schon heute mit dem Aufbau neuer Kernkompetenzen begonnen werden, um auf die sich ständig verändernde Medienwelt reagieren zu können.

Auch wenn derzeit fast jede elektronisch publizierte literarische Edition parallel zu der zugehörigen Buchausgabe erscheint, muss dies nicht so bleiben. Hat ein Verlag die Entscheidung für einen Strategiewechsel einmal getroffen, so macht er sich auf einen neuen Weg, der schwierig ist, da noch auf wenige Erfahrungen anderer Verlage zurückgegriffen werden kann. Eine solche Neuorientierung trägt der veränderten Medienlandschaft Rechnung, muss aber, wenn sie kein modisches Mitmachen sein soll, auch mit einem strategischen Wollen einhergehen, das nicht nur das Morgen, sondern auch das Übermorgen mit einbezieht. Solche Richtungswechsel werden bei einem Verlag nicht im luftleeren Raum entschieden; sie gehen in aller Regel mit konkreten Projektplänen einher.

So hat der S. Fischer Verlag das Projekt der 'Großen kommentierten Frankfurter Ausgabe der Werke Thomas Manns' zum Anlass genommen, ein zukunftsweisendes verlegerisches Konzept zu erstellen und dabei neue Strategien im Umgang mit Inhalten zu etablieren. In diesem Rahmen wurde ein Informationspool konzipiert, der das verlegerische Herzstück für das Werk Thomas Manns darstellt, und ein semantisches Netz, das einen vollkommen neuen Umgang damit ermöglicht. Mit dem semantischen Netz gelingt zusätzlich eine differenzierte redaktionelle Erschließung der Inhalte, wodurch es die aktuelle verlegerische Aufgabe erfüllt, diese Inhalte für verschiedene Publikationsmedien zu öffnen. Dies ermöglicht eine neuartige Organisation der Ausgabe und bedeutet in der Konsequenz eine Zukunftssicherung des Thomas-Mann-Werks.

[Kindlers neues Literaturlexikon. 21 Bände auf einer CD-ROM. München 1999] zugrunde, so sind semi-automatische Methoden denkbar.

Der S. Fischer Verlag stellt nach wie vor schöne Bücher her. Er hat sich aber auch auf den Weg gemacht, seine Inhalte so aufzubereiten, dass er für die neuen Medien gerüstet ist. Die neuen Publikationsmärkte für CD-ROM, E-Book und Internet verändern sich beständig. Es ist heute mehr denn je notwendig, auf die Veränderung dieser Märkte flexibel reagieren zu können und sie mit qualitativ hochwertigen Produkten zu bedienen.

Dieser neue Umgang mit den Inhalten geht einher mit Veränderungen im Verlag: Neue Aufgaben entstehen, schon bestehende Aufgabenbereiche verzahnen sich stärker. Es ist wichtig, diese Veränderungen als Chance und Herausforderung zu begreifen.